POLICY RESEARCH WORKING PAPER 3112

On the Timing of Marriage, Cattle, and Weather Shocks in Rural Zimbabwe

Johannes Hoogeveen
Bas van der Klaauw
Gijsbert van Lomwel

The World Bank
Development Research Group
Poverty Team
August 2003

Abstract

Hoogeveen, van der Klaauw, and van Lomwel focus on the timing of marriages of women in rural Zimbabwe. Zimbabwean marriages are associated with bride wealth payments, which are transfers from (the family of) the groom to the bride's family. Unmarried daughters could therefore be considered assets who, at time of need, can be cashed in. The authors investigate to what extent the timing of a marriage of a daughter is affected by the economic conditions of the household from which she originates. They distinguish household-specific wealth levels and two types of shocks—correlated (weather) shocks and idiosyncratic shocks. The authors estimate a duration model using a unique panel survey of Zimbabwean smallholder farmers. The estimation results support the hypothesis that the timing of marriage is affected by household characteristics. Girls from households that experienced a negative (idiosyncratic) shock in their assets are more likely to marry.

This paper—a product of the Poverty Team, Development Research Group—is part of a larger effort in the group to assess how informal institutions respond to the needs of the poor. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Patricia Sader, room MC3-632, telephone 202-473-3902, fax 202-522-1153, email address psader@worldbank.org. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. Johannes Hoogeveen may be contacted at jhoogeveen@worldbank.org. August 2003. (45 pages)

On the Timing of Marriage, Cattle and Weather Shocks in Rural Zimbabwe

Hans Hoogeveen *
Bas van der Klaauw †
Gijsbert van Lomwel ‡

*World Bank and Free University Amsterdam.
Address: The World Bank, DECRG MSN MC3-305, 1818 H Street N.W., Washington D.C. 20433, USA.
†Free University Amsterdam and Scholar.
Address: Department of Economics, Free University Amsterdam, De Boelelaan 1105, NL–1081 HV Amsterdam, The Netherlands.
‡CentER Applied Research.
Address: Tilburg University, P.O. Box 90153, NL–5000 LE Tilburg, The Netherlands.

Keywords: coping strategies, bride wealth, duration models, initial conditions, wealth shocks.

We are grateful to Bill Kinsey for providing the data. We thank Marleen Dekker, Paul Frijters, Hanan Jacoby, Geert Ridder, Mark Rosenzweig, Chris Udry, seminar participants at the World Bank, Tinbergen Institute Amsterdam, and Tilburg University, and the participants of the 2001 meeting of the European Economic Association in Lausanne for useful comments.

1 Introduction

In developing countries economic markets are often absent or incomplete. Zimbabwe is no exception. The absence of economic markets is particularly severe in rural areas, where smallholder farmers lack access to formal financial and insurance institutions. This implies that these farmers often face liquidity constraints and that they cannot insure against a bad harvest or loss of livestock. Whereas a bad harvest has immediate but generally temporary consequences for farmers' welfare, the consequences for welfare of loss of livestock are long lasting. In rural Zimbabwe income from agricultural production mainly depends on two factors, sufficient rainfall and the availability of cattle. Cattle are necessary as draft power for plowing. Without cattle households can only produce using the hoe, which yields little income. Loss of livestock might therefore lead to a household being stuck in poverty for a considerable period of time (see Carter and Zimmerman, 2000).

The lack of formal financial institutions does not allow farmers to borrow for investments in agricultural production or buying cattle. A large range of non-market solutions has become available to insure households against negative shocks and to allow them to smooth consumption over time. Examples of such non-market solutions are food sharing arrangements in which a household, who faced a negative idiosyncratic shock gets food from other households in the village, and labor sharing arrangements under which within a village farmers take care of the land of someone who is temporary unable to do so. Rosenzweig (2001) provides a review of the literature on informal insurance arrangements in low-income countries.

The focus of this paper is on the timing of marriage as an alternative institution for insurance against the loss of livestock. That marriage may act as a non-market insurance is due to the fact that marriage in Zimbabwe involves bride wealth payments. These payments are made by (the family of) the groom to the bride's family. Bride wealth comprises of a substantial number of cattle of which a considerable fraction is paid at the moment of the actual marriage. An unmarried daughter thus represents access to livestock and her marriage may be considered an asset that can be cashed in times of adversity. Because of bride wealth payments, marriages should be considered as a contract between two families rather than between two individuals. However, the choice of the spouse is generally not a household decision, but an individual decision. Therefore, we do not focus on the choice of spouses, but instead consider the timing of the marriage.

The bride's family can use the cattle obtained as bride wealth for increasing agricultural production and as buffer stock. The marriage decision of a daughter is therefore more likely to be a household decision in poor households than in wealthier households. In the empirical analyses, we will investigate to what extent the timing of marriage of young women can be explained by the economic conditions of the households from which they originate.

There are many aspects to how shocks affect marriages. It is particularly important to distinguish between correlated shocks induced by a lack of sufficient rainfall and idiosyncratic shocks to household for example as a result of theft or death of cattle. Unlike idiosyncratic shocks, correlated shocks may have equilibrium effects on the marriage market. If after a year of low rainfall households press daughters to marry, households may prefer that sons do not marry. It may therefore be the case that after a year of low rainfall actually less marriages occur. However, this may also have consequences for the amount of bride wealth. If after a year of low rainfall women are more eager to marry than men, then the amount of bride wealth is expected to be lower. We return to these issues in the empirical analyses, where we will distinguish between correlated and idiosyncratic shocks and where we also consider the amount of bride wealth.

A strong impact of the household's economic situation on the timing of marriage suggests that formal risk sharing and credit markets work indeed unsatisfactorily. Such knowledge can be beneficial for policy, especially in the context of economic development. Furthermore, the presence of the use of bride wealth as an informal insurance device may cause women to marry younger. This may affect fertility and might increase the average household size. This thus has implications for population growth.

A dynamic model is required to empirically analyze the timing of a marriage. Such a model is required because the composition of the population of unmarried women in a particular year depends on marriages in previous years. If due to some shock, for example a drought, in a particular year many women in poor households marry, it follows that in the following year the population of unmarried women shifts towards women from richer households. To account for this we use hazard rates to model the age of marriage. We allow the transition rate from being unmarried to being married (the marriage rate) to depend on observed explanatory variables, both individual, household and environmental characteristics, on the elapsed duration of being unmarried (age), and on unobserved determinants. To model the relationship between the marriage rate and these variables, we use a Mixed Proportional Hazard specification (see e.g. Lancaster, 1990). To estimate the model we use subsamples of the data constructed under two different sampling schemes, namely flow and stock sampling. The use of a stock sample implies that we have to correct for initial conditions, while constructing a flow sample reduces the number of observations. We also distinguish different shocks that occur to households. To construct idiosyncratic shocks, we use a dynamic model for household wealth accumulation. This allows us to separate idiosyncratic shocks from household specific wealth levels and correlated shocks. Finally, we also perform empirical analyses on the amount of bride wealth payments. However, due to the limited information on the amount of bride wealth involved with the observed marriages, the empirical results based on the size of bride wealth should be interpreted with care.

The data we use are from an annual panel survey held under a group of resettlement farmers in Zimbabwe which started in 1982. In the earlier surveys only concerned household level variables, from 1994 onwards also detailed information on all members has been collected. We use the subset of the data that starts in 1994, which includes approximately 400 households. The data on household members allows us to construct subsamples of unmarried daughters, which include information on the marriage decision, individual characteristics and some measures of household wealth. Since the data we use in the empirical analyses deal with farmers, in the remainder of the paper we restrict our discussions to farmers and rural areas.

The paper is organized as follows. In Section 2 the institutions of marriage in rural Zimbabwe are considered in more detail. Section 3 presents the statistical model. Section 4 describes the data. The estimation results are presented in Section 5. Section 6 concludes.

2 Cattle and marriage in Zimbabwe

In the introduction the importance of cattle to Zimbabwean smallholder farmers as source of draft power for transportation and especially for plowing was already stressed. The heavy loam soils in rural Zimbabwe require at least two head of cattle, preferably trained oxes, to prepare land for sowing. Without two draft animals, a household has little choice but to rely on labor intensive land preparation using the hoe. The income obtained from farming manually is low, and insufficient to buy new cattle. Money can be obtained in off farm jobs, but these jobs are hard to find (during our observation period the urban unemployment rate was around 50%). Because of the absence of formal financial markets, borrowing money to buy cattle is impossible. A household without cattle can try to borrow cattle from neighbors after they are done with plowing. Typically, by the time the animals are available, the time for planting has passed while the beasts are tired from their earlier effort and of little use. Households without cattle can try to save (in the form of goats and chicken) from their meager crop incomes, and in this way accumulate the wealth to buy cattle. Yet it might take a long time before a sufficient amount is accumulated (Carter and Zimmerman, 2000). Households therefore have a strong incentive to avoid that the number of cattle becomes too small.

There are other reasons why cattle are of great significance to Zimbabwean farmers. Cattle provide milk, produce manure to fertilize the land and in times of need they can be eaten or sold in return for grain.[1] Selling cows to buy food is only relied upon as a measure of last resort. Fafchamps, Udry and Czukas (1998) show for Burkina Faso for instance that during long periods of drought, cattle sales only compensated 15% to 30% of income losses. Still, in Zimbabwe during the 1992 and 1995 droughts, the sale of cattle and other livestock was the single most important way to obtain cash to purchase food (see Kinsey, Burger and Gunning, 1998). The absence of formal financial institutions not only makes cattle an important buffer stock it is also an important means to keep savings (e.g. Fafchamps, Udry and Czukas, 1998; Rosenzweig and Wolpin, 1993). In this light, weight increase of a cow and the birth of a calf can be considered as interest on these savings, while the aging of the cattle represents inflation. Since the possession of cattle is vital for agricultural production, cattle ownership in Zimbabwe is highly correlated with other forms of asset ownership (Scoones, 1995). Furthermore, cattle can be used as legal tender. Not only to settle debt, but also in civil or criminal cases a traditional court may convict someone to pay a number of cattle. Finally, Tsodzo (1992) argues that also a great social value is attached to possessing many head of cattle ("he who dispenses with them,

[1] Fafchamps, Udry and Czukas (1998) note that African farmers rarely kill their cattle for eating the meat. Zimbabwean farmers may be even more reluctant to do so, since unlike in Burkina Faso, cattle are used for agricultural production.

dispenses with good living").

Next let us turn to the process of marriage in Zimbabwe. It is characterized by duality as the choice of one's spouse is left to the individuals concerned, while the marriage itself is considered a contract between families. The latter is illustrated by the relationship terminology that is adopted. After marriage, the groom's father-in-law becomes father-in-law to the whole of the groom's family, while the father of the groom is regarded as the principle son-in-law rather than the groom himself. A brother of the groom is likely to speak in the first person saying "I have married such and such family".

A marriage provides a bond between two families, and Zimbabwean families related through marriage typically share resources in an effort to deal with risks. Families prefer to spread the net of affinal relations by marrying into different families. Marriages over a long distance or with someone with a job in town are considered with favor as these mitigate the impact of local weather shocks (see also Rosenzweig, 1988, and Rosenzweig and Stark, 1989).

Zimbabwean marriages include bride wealth payments, which are transfers from the family of the groom to the family of the bride. The direction of the payments depends on the relative scarcity of land compared to labor. In the relatively scarcely populated rural areas of Zimbabwe, the groom purchases the production capacity of the bride and of future children (e.g. Jacoby, 1995). In India, where land is scarce compared to labor the payments are usually made in the opposite direction, which is referred to as dowry (e.g. Rao, 1993).[2]

Bride wealth in Zimbabwe consists of two distinct payments, which are referred to as *rutsambo* and *danga*. *Rutsambo* is associated with sexual rights to the woman and after its payment the girl is allowed to move to her husband's household. *Danga* is associated with rights over children born to the woman. Cattle are the main body of the bride wealth. On average 9 head of cattle are demanded as bride wealth.

At the moment of marriage a substantial part of *danga* is paid, but the bulk of *danga* remains outstanding. Full payment of *danga* is extended over a long period of time. In Figure 1 we show how the fraction of marriages where all bride wealth payments are made evolves over the duration of the marriage. For less than 10% of the marriages all bride wealth payments are made in the first few

[2]See Binswanger and McIntire (1987) and Binswanger and Rosenzweig (1986) for extensive discussions on the consequences of geographic characteristics on economic institutions.

years after the marriage. For over 80% of the marriages some part of the bride wealth payments is still outstanding after 25 years of marriage. The large drop occurs around 30 years of marriage, which implies that for most marriages the final bride wealth payments are made between 28 to 35 years of marriage. The bulk of the bride wealth thus remains outstanding after the marriage, which has advantages for both the family of the bride and the family of the groom. For the son-in-law, delayed payment implies that he can pay when he has the means to do so. Moreover, he can make sure that his wife is childbearing. Being barren is a valuable reason to undo a marriage. After a divorce or if the wife dies without giving birth to many children the husband can claim back (part of) the bride wealth. On the other hand, as long as the bride wealth is not completely paid, the family of the bride can ask the groom for favors in the form of services and gifts. This is illustrated by a Zimbabwean proverb saying: "A son-in-law is like a fruit tree: one never finishes eating from it". Furthermore, delayed payments secure that a household that loses its cattle still has access to some assets (see Dekker and Hoogeveen, 2002). Therefore, even if the son-in-law is in the position to pay all bride wealth upon marriage, it is considered a denial of the marriage bond between the families involved to actually do so.

The optimal timing of marriage is evident. Unless the household is very wealthy sons should marry late and daughters early. Late marriage of a son has several advantages. The son remains productive in the family's agricultural activities and the loss of draft power is postponed. This allows (the family of) the groom more time to accumulate cattle. If a household has many head of cattle, the marginal productivity of cattle decreases, as there typically is no labor market and the amount of family labor and land are fixed. So, if a family has sufficient cattle, the marriage of a son is relatively rewarding in terms of additional labor received (the wife) and relatively cheap in terms of cattle lost. If a household is poor, daughters should marry early. After the marriage of a daughter the previously poor family is able to experience a period of high productivity before the marriage of a son is supported. This period of high productivity can be used to accumulate more cattle and to grow out of the initial period of poverty. Another reason why daughters should marry young is that the amount of bride wealth decreases with age, which reflects that relative to older women, young women are likely to give birth to more children. National data confirm this pattern; the marriage age of men is much higher than of women (see Central Statistical Office, 1995). The median age at first marriage for men is 25 years, compared with 19 years for women. Only 11 percent of the men are married by the age of 20, compared to 62 percent of the women. The optimal timing of marriage provides a testable implication. In the next sections we investigate empirically whether it is the case that daughters from poorer households marry younger than daughters from wealthier households.

So far no attention has been paid to the risks faced by the households. Cattle are subject to several risks such as plagues like pests, theft, and periods of drought. Pests have had devastating effects in the past, but are largely brought under control. Theft is a regular phenomenon, and can cause serious problems for the household. The consequences of weather shocks are substantial as well. Scoones, Chibudu, Chikura, Jeranyama, Machaka, Machanja, Mavedzenge, Mombeshora, Mudhara, Mudziwo, Murimbarimba and Zirereza (1996) show that in the Chivi region (in Southern Zimbabwe) after the 1982-1984 droughts the number of households without cattle more than doubled from 23.3% to 50.7%. During a period of drought, farmers sell some cattle to buy food necessary to survive (see Kinsey, Burger and Gunning, 1998). Yet because recovering from a period of drought is almost impossible without any cattle, they are very reluctant to sell cattle. In Table 1 we present cattle survival rates for the years 1987/1988 and 1991/1992. The year 1987/1988 was a year with relatively abundant rainfall (744mm), while the year 1991/1992 was a year of drought (only 335mm rainfall). The table illustrates that in the period of drought herd losses of more than 50% were not exceptional.

Formal institutions that allow smallholder farmers to insure against these risks do not exist. Therefore, households have to use alternative ways to deal with negative shocks. Households can increase their wealth as precaution against negative shocks. Above we already argued that cattle are the main assets of a household. Increasing the herd size might be an option to anticipate the consequences of negative shocks. Alternatively, postponing the marriage of a daughter can be considered as a substitute for precautionary saving. And since unmarried daughters are not subject to the risk of theft, postponing the marriage of a daughter can be considered a portfolio diversification. After a negative shock the household can exercise the 'option value' of an unmarried daughter to raise the herd size again. The testable implication is that shocks affect the marriage behavior of women. After a negative shock households will press their daughters to marry. In the next sections we will empirically analyze the relation between shocks and the age of marriage. We will distinguish between idiosyncratic shocks, such as theft and death of cattle, and correlated shocks, such as a period of drought. The latter shocks might also have equilibrium effects, since after a correlated negative shock less households are willing to pay bride wealth for sons getting married.

3 Specification of the statistical model

In this section we present our empirical model. We focus on unmarried women and the transition to being married, the marriage rate. We allow the marriage rate to depend on age, calendar time and a number of household characteristics. The empirical model is based on commonly used hazard rate models (see Lancaster, 1990). First, we discuss the specification of the marriage rate. Next, in Subsection 3.1 and Subsection 3.2 we discuss two sampling schemes used to obtain data sets, stock sampling and flow sampling. We mainly focus on the consequences of the sampling scheme for the restrictions necessary to derive proper loglikelihood functions. We conclude this section with the parameterization of the model.

It is useful to start with a brief outline of our data and to provide some characteristics of the data that are relevant for the model. A more elaborate discussion of the data is given in the next section. The database includes unmarried women, who are interviewed yearly. At the moment of the interview the women are asked about their marital status and age. Additionally, a number of household characteristics are collected. This allows us to construct spells of being unmarried, although we do not observe the exact date of marriage and the date of birth, but only the years. Some women are still unmarried at the end of the observation period, and some other women left the database unmarried, for example because they temporarily moved to relatives. Such spells of being unmarried are right-censored, which we treat as exogenous.

It is clear that it is not necessary to observe women from the moment of birth: it is sufficient to follow them from the age they start marrying. We set the minimum age of marrying at 15, which is denoted by t_0. This minimum age is not due to legal restrictions, but reflects the age the statistical bureau of Zimbabwe uses as lower bound for the age at marriage. Of the women born between 1975 and 1980 less than 3 percent was reported to be married by the age of 15 (Central Statistical Office, 1995). This is also confirmed by our data, which does not show any marriages of women under age 15.

Age and calendar time are genuinely continuous, but we only observe in which year a woman got married. To deal with this discrepancy, we use an underlying continuous-time marriage rate to model the probability that an unmarried woman of age $t \geq t_0$ at calendar time τ marries before calendar time $\tau+1$. Since we do not observe the exact date of birth, but only the age at the moment of the interview, we integrate over the possible dates of birth. We assume that the differences in transition rate from unmarried to married can be characterized by calendar time τ, a vector of observed individual and household characteristics x, unobserved individual characteristics v and the elapsed duration of being unmarried t (age of the unmarried woman). We assume v to be independent of x and τ. The unobserved characteristics are thus individual specific and not household specific. The covariates x are not necessarily constant over time, however changes over time in x are only observed to occur at the moment of the interview and are assumed to be exogenous and unpredictable. We suppress subscripts explicitly denoting that x is time-varying.

The marriage rate at age t conditional on x, τ and v is denoted by $\lambda(t|\tau, x, v)$ and is assumed to have the familiar Mixed Proportional Hazard (MPH) specification

$$\lambda(t|\tau, x, v) = \psi_1(t)\psi_2(\tau)\exp(x\beta + v),$$

in which $\psi_1(t)$ and $\psi_2(\tau)$ are positive functions representing the baseline hazard or age dependence and calendar time effects, respectively. The calendar time effects $\psi_2(\tau)$ are considered to be changes in economic conditions faced by the household, which will be represented by changes in rainfall. Let t be the actual age when getting married and τ_0 the calendar time at birth. The conditional density function of $t|\tau_0, x, v$ can be written as

$$f(t|\tau_0, x, v) = \lambda(t|\tau_0 + t, x, v)\exp\left(-\int_{t_0}^{t} \lambda(s|\tau_0 + s, x, v)\mathrm{d}s\right), \qquad t \geq t_0.$$

The conditional survivor function of $t|\tau_0, x, v$, i.e. the conditional probability that the duration of being unmarried exceeds t, equals

$$S(t|\tau_0, x, v) = \exp\left(-\int_{t_0}^{t} \lambda(s|\tau_0 + s, x, v)\mathrm{d}s\right), \qquad t \geq t_0.$$

Obviously, $S(t|\tau_0, x, v) = 1$ if $t < t_0$, this holds for any value of τ_0, x and v.

Note that in the model we treat each woman as an independent observation. This is not completely correct as some women live as sisters in the same household. For a household with only few head of cattle, one of the daughters getting married can be sufficient to avoid future poverty, implying dependency between the marriage rates of sisters. In particular the marriage rates between sisters are likely to be negatively correlated. It also implies that by treating sisters as independent, we underestimate the effect of household's economic conditions on the marriage rate. We do not correct for this. The parameters we estimate are no policy parameters and therefore we are more interested in their signs rather than their exact values.

In case we observe women from the moment they reach age t_0 inference is straightforward. However, in a given year the data are sampled from the stock of unmarried women. In the following years we observe whether a woman gets married or remains unmarried. At this point there are two possible ways to proceed. We can either construct a flow sample by only considering women reaching age 15 during the observation period or we can use the complete (stock) sample of unmarried women over age 15. The main disadvantage of the first approach is that a large share of the data is ignored. On the other hand the use of a stock sample of unmarried women causes initial condition problems, as some women who were older than 15 and unmarried at the beginning of the observation period had already been exposed to the hazard of getting married for some years (see Ridder, 1984). To solve these initial condition problems it is necessary to make some additional assumptions on the marriage process, which we discuss later in more detail. Both sampling schemes therefore have disadvantages, and we decide to use both. We estimate the model first using the stock sample of unmarried women and second using only a flow sample of women reaching age 15.

3.1 Inference based on stock sampling

When using the full (stock) sample, the initial condition problems mentioned earlier arise because the age of the women at the moment of entering the database differs. In the economic literature some solutions to initial condition problems in the context of duration models have been proposed. For example, Flinn and Heckman (1982) suggest to specify a separate hazard function for spells already in progress at the beginning of the observation period. This approach can only be applied if the data contain multiple spells for single individuals, which is obviously not the case when modelling the age at which a woman first marries. The solution we use is from Ridder (1984), who derives explicit expressions for the distribution of left-censored spells, i.e. the distribution of the age of marriage of women who were older than 15 when they were first observed in the data. This implies that inference is based on the joint distribution of the age at entering the sample and the age of marriage, conditional on actually being in the sample. This method does not require multiple spells, but it requires additional stationarity assumptions. First, we assume that the rate at which women reach age 15 has been constant in the years before the observation period, i.e. each time period the same number of women reach age 15. Second, we need to assume that the composition of women reaching age 15 in terms of observed and unobserved characteristics has been constant over time. And third, we can only correct for time-varying regressors if the complete history of such a regressor is observed and the regressor affects the marriage rate of all women in the same way. The only time-varying regressor for which this holds is the amount of rainfall, which we also observe before the observation period. Other possibly important explanatory variables such as the household wealth are not known for the period before the observation period and thus cannot be included in the marriage rate. Therefore, the third assumption may be too restrictive, as from the discussion in Section 2 we know that the timing of marriage is a decision of both the household and the individual, in which the household's pressure depends on current wealth.

Let $n(\tau)$ be the rate at which women reach age t_0 at calendar time τ. An indicator of $n(\tau)$ could be the birth rate in year $\tau - t_0$. We assume that the composition of the group of women reaching t_0 does not change over calendar time and also that household characteristics are constant over time. At the date of sampling τ_s, the total stock of unmarried women with observed characteristics x is proportional to

$$N(\tau_s|x) \propto \int_{t_0}^{\infty} n(\tau_s - t + t_0) \int_v S(t|\tau_s - t, x, v)\mathrm{d}G(v)\mathrm{d}t,$$

in which $S(t|\tau_s - t, x, v)$ is the survivor function, i.e. the probability that a woman with characteristics x and v, who was born at date $\tau_s - t$ is still unmarried at age t, and $G(v)$ is the distribution function of the unobserved characteristic v.

Suppose in a given year τ_s we observe a sample of unmarried women. For each woman we observe her age t and a vector of individual and household characteristics x. Furthermore we observe that this woman gets married between year τ^* and $\tau^* + 1$. Since we do not observe the exact day at which a woman reaches t_0, we have to integrate over the year at which this occurred. The same holds for the moment of marriage, which is also only observed at yearly intervals. Let R denote the residual duration of being unmarried after the date of sampling and P the past duration of being unmarried before the date of sampling. The probability of this observation equals

$$\Pr(R \in [\tau^* - \tau_s, \tau^* - \tau_s + 1), P \in [t, t+1) |\text{ In sample at } \tau_s, x)$$
$$= \frac{\int_{\tau^*-\tau_s}^{\tau^*-\tau_s+1} \int_t^{t+1} n(\tau_s - p + t_0) \int_v f(p + r|\tau_s - p, x, v) \mathrm{d}G(v) \mathrm{d}p \mathrm{d}r}{N(\tau_s|x)}$$

The individual contributions to the loglikelihood function are equal to the logarithm of this probability.[3] Within this hazard rate framework right-censoring, i.e. a woman is still unmarried at the end of the observation period, is solved in a straightforward manner.

3.2 Inference based on flow sampling

A flow sample is not subject to initial condition problems as it only contains women who are observed from the moment they reach the age at which they start getting married. Since the flow sample only includes women reaching age 15 during the observation period, the cost of avoiding initial condition problems is ignoring a large share of the available data.

Due to the limited length of the observation period we are not able to specify a complete path of the baseline hazard. On the other hand we do not have to make assumptions on the stationarity of the marriage rate, i.e. the use of a flow sample allows us to include time-varying explanatory variables. The last difference with the stock sampling approach is that we do not need any information on the rate at which women reach age t_0 (which was denoted by $n(\tau)$).

As in the stock sampling case, using the flow sample we have to integrate over yearly intervals in which the women reached age 15 and got married. Let T be the stochastic variable denoting the date at which a woman gets married. Conditional on the year of inflow in the sample τ_s, i.e. the year in which the women became 15 and the vectors of explanatory variables x_τ, $\tau = \tau_s, \tau_s + 1, \ldots$,

[3] It is also possible to base inference on the conditional probability

$$\Pr(R \in [\tau^* - \tau_s, \tau^* - \tau_s + 1) | P \in [t, t+1), \text{ In sample at } \tau_s, x)$$

This provides consistent but less efficient estimators (see Ridder, 1984).

the probability of getting married between year τ^* and τ^*+1 equals

$$\begin{aligned}&\Pr(T \in [\tau^*, \tau^*+1)|\text{ Inflow at } \tau_s, x_{\tau_s}, \dots, x_{\tau^*})\\&= \int_{\tau^*-\tau_s}^{\tau^*-\tau_s+1} \int_{t_0}^{t_0+1} \int_v f(\tau + p|\tau_s - p, x_{\tau_s}, \dots, x_{\tau^*}, v)\mathrm{d}G(v)\mathrm{d}p\mathrm{d}r\end{aligned}$$

The individual contributions to the loglikelihood function are based on this distribution function and like in the stock sampling approach right-censoring is solved in a straightforward way.

3.3 Parameterization

For the baseline hazard function, $\psi_1(t)$, and the unobserved heterogeneity distribution, $G(v)$, we take the most flexible specifications used to date. We take $\psi_1(t)$ to have a piecewise constant specification,

$$\psi_1(t) = \exp\left(\sum_{j=1,2,\dots} \lambda_j I_j(t)\right),$$

where j is a subscript denoting time intervals and $I_j(t)$ are time-varying dummy variables that are equal to one in consecutive time intervals. Note that with an increasing number of time intervals any pattern can be approximated arbitrarily closely. We take these time intervals equal to two years.

We take the distribution of the unobserved heterogeneity $G(v)$ to be a discrete distribution with a fixed number of unrestricted mass point locations. Let v_i, $i = 1, \dots, m$ be the points of support and let p_i, $i = 1, \dots, m$ be the associated probabilities, i.e. $p_i = \Pr(v = v_i)$, where $0 \leq p_i \leq 1$, $i, = 1, \dots, m$ and $p_m = 1 - p_1 - \dots - p_{m-1}$. Note that discrete distributions are advantageous not only from the point of flexibility but also from a computational point of view.

We also take the calendar time effects $\psi_2(\tau)$ piecewise constant. The calendar time effect in a given year is a function of the quantity of rainfall in the previous agricultural season, which runs from the fourth quarter of a year to the first quarter of the next year. In Zimbabwe most marriages take place during the fourth quarter of a year. It is therefore most likely that the marriage decision in year t is affected by the quantity of rainfall in the harvest season during the fourth quarter of year $t-1$ and the first quarter of year t.

4 Data

The data set we use is a yearly panel of Zimbabwean smallholder farmers covering the period 1994 until 2000.[4] These households belong to the group of about 25,000 families that have been resettled in the early 1980s on land acquired from large-scale farmers after the independence. Approximately 400 households are interviewed, living in three different regions (Mpfurudzi, Sengezi and Mutanda). These schemes were chosen to ensure representation of each of the three major agro-ecological zones in the country which are suited for cropping. Farmers located in Mpfurudzi live in the area most favorable to farming, those in Mutanda have to deal with the worst conditions. Those resettled in Sengezi live in an intermediate area.

Upon settlement each household was provided with 12 acres of land. The land presented to the households is about 10% of the total area; the remaining 90% is common property land, which might be used for grazing. In the data set 79% of the households uses less than the 12 acres of land for agricultural production, 9% uses exactly 12 acres and the remaining 12% uses more than 12 acres. For most households the availability of land is thus not the binding constraint in their agricultural production especially as households can use more than 12 acres of land if they (illegally) reclaim some common property land or rent land from other farmers.

Farmers in resettlement schemes differ from regular smallholder farmers in Zimbabwe. The resettlement farmers possess more land and until 1992 heads of households were not allowed to work off farm. This is reflected in the way income is generated. Unlike regular farmers who obtain approximately 30% of their income from non-farm sources, the resettlement farmers earn only 5% to 10% of their income off farm. Farmers in the resettlement scheme are also somewhat wealthier than regular farmers and their per capita expenditures are about 10% higher than the per capita expenditures of ordinary smallholder farmers (Deininger and Hoogeveen, 2002).

The data are collected by Bill Kinsey and are described in more detail in Kinsey, Burger and Gunning (1998). The interviews take place in the beginning of the year, and thus describe the previous year. The rainy season is from the

[4]The data also includes the survey for 2001. However, we decided not to use this information because starting in 2000 large changes took place in Zimbabwe, especially in the political situation regarding land reform.

end of November until the beginning of April. This is immediately followed by the harvest period which takes until June/July. Marriages usually take place in the fall. Starting in 1994 information on all household members is collected along with household level variables, whereas the 1993 interview only collected household level information.

The survey contains both questions about the household and about all individuals in the household. At the household level variables such as crop income and livestock possession are administered. The questions to the individuals are about age, level of education, gender, marital status, etc. If an individual is not present at the time of the survey, but was recorded as a household member in a previous survey, the reason is asked why the individual left the household. Women generally leave the household upon marriage. So by comparing presence in the household and marital status in consecutive years, it is possible to construct spells of being unmarried.[5]

Since we are interested in the marriage behavior of young women, we only use the subset of the data consisting of all unmarried women whose age exceeds 15 years at some interview. Furthermore, we restrict the subsample to women under the age 30, as older unmarried women are often divorced or widowed. In total we observe 691 unmarried women over the age 15, of which in total 233 got married during the observation period. We refer to this sample as the stock sample. The empirical marriage rate is shown in Figure 2. It reflects an increasing pattern until age 21. After that the empirical marriage rate fluctuates, but there is no clear trend. Figure 3 shows the corresponding Kaplan-Meier estimates of the survival probabilities, i.e. the probability that a woman is still unmarried at a particular age. Most women marry before age 30, only 4% is still unmarried at this age. At age 22 over 60% of the women has got married. The median age of marrying in our data is around 21 years and thus higher than the median marriage age in Zimbabwe, which is 19 years (see Section 2). Remember that our data describe a very specific rural area, where the (resettlement) farmers are on average somewhat richer than other farmers.

In the stock sample around 66% of the spells are right-censored. Right-censoring arises for two reasons. First, a woman can still be unmarried at the end of the observation period. Second, a woman can leave the household (therefore

[5]Generally daughters do not leave the household for reasons other than marriage. Sons are much more likely to leave the household unmarried to seek employment in town and get married while they are in town. This makes it more difficult to focus on the marriage behavior of sons.

disappearing from our data set) without having got married. This happens if a woman stays for some time with family or relatives in another area, moves to work elsewhere, or in case the woman dies.

In Table 2 we present some annual statistics of the data, such as the number of marriages, the average marital age and both local and average rainfall in Zimbabwe.[6] We have stratified the sample by the region in which the women live. For each of the three regions both the number of marriages and the average age of marriage decrease over the years. Also the sample size decreases over the years. This is caused by the original sampling of the data, which includes households of which the man was between 35 and 50 years old in 1982.[7] This implies that more women leave the sample (for reason of marriage), than young (unmarried) women enter the sample (because they reach age 15). Local rainfall is correlated with average rainfall in Zimbabwe, but there is variation between the regions. In each of the regions rainfall was lowest in 1994/1995, which affects the marriage in the Fall of 1995. The data do not show strong differences in the number of marriages and the average age of marriage between the years with a large and low amount of rainfall.

Table 3 provides the average marriage age stratified by the year of marriage and the herd size of the household. In most years, daughters in the less wealthy families marry younger than in the wealthier families. However, the differences are not very pronounced.

As explained in Subsection 3.1, the empirical analysis based on the stock sample only allows for stationary explanatory variables. Not many variables are stationary. In Table 4 we give the characteristics of the stock sample. These characteristics are computed at the individual level (of unmarried women), and should thus not be compared with averages at the household level. Most women in the sample live in villages in Mpfurudzi. In terms of religious denomination around 17% of the women reports to have African faith and 6% Masowe faith. The remaining women mostly have Christian faith, which is generally less strict than the African and Masowe faith. For around 10% of the women household characteristics are missing, we include a dummy variable for these women. Household size is not really a stationary variable. For the moment we take it as being

[6]The local rainfall data and the Zimbabwean rainfall data are collected by the Department of Meteorological Services in Zimbabwe.

[7]A family remains in the data until both the man and woman die. After that this household is replaced by the people who move into their house (most likely sons or daughters).

stationary, measuring the household size at the moment a girl first appears in the stock sample. Household size is relatively constant over the years, on average the women live in households containing around 12 persons (including the parents). In 17% of the cases the father is missing in the household and in 4% the mother. The father is usually older than the mother and has on average 1 year of education more. These are the only variables in the database that are stationary.

The flow sample consists of all women who reached age 15 during the observation period. This gives a sample of 333 women (originating from 302 different households), of which 57 got married during the observation period. Table 5 shows the number of women in the flow sample at the beginning of a year and the number of marriages in that year. Because the sample size is relatively small we do not stratify by region. It is clear from the table that most marriages occur in the last three years of the observation period. During the first three years, the women in the flow sample are all younger than 18 years. The marriage rate for this age group is low (as can be seen from the empirical marriage rate in Figure 2).

In Section 2 we stressed the relation between household wealth and marriage behavior. We argued that cattle is the most important asset of households. As a measure of household wealth we therefore include livestock wealth expressed in the number of cows. In Table 6 we show the average livestock wealth in each year for the subsample of women who got married in that year and those who remained unmarried. We see that in each year except for 1999, the women who got married came from poorer households. Recall that in the first three years the number of marriages was very low. The difference in wealth between the subsamples is particularly large in 1998, which was a year following the relative dry year 1997. It seems a confirmation of the household insurance hypothesis suggesting that after a negative shock women in poor households are more likely to get married. However, it might be the case that there are households that are always poor and households that are always rich. If girls from poorer households prefer to marry younger for instance because living in a poor household is not very pleasant, then we could observe a similar pattern. To get an indication on how big shocks in household wealth are compared to the distribution of wealth between the households, we compare the average standard deviation of livestock wealth of households (over the years) with the standard deviation of average wealth of households (over the years). The average value of livestock is around 11.7 cows, the average standard deviation of household livestock wealth 3.3 cows and the standard deviation of average wealth of households 8.9 cows. This implies that there are indeed generally richer and poorer households, but being rich in one particular year does not guarantee being rich in the next year. In general, we see that the average livestock wealth of households is increasing during the observation period as households are recovering from the drought in 1991/1992. We return to the issue of separating shocks in livestock wealth from household specific livestock wealth levels in the next section. Another variable we include is an indicator for being the oldest unmarried daughter in the household. If the timing of a marriage is indeed a household decision, a (temporarily) poor household will most likely put pressure on the oldest unmarried daughter to get married. Furthermore, in the flow sample we treat household size as a time-varying regressor.

The data contain some information on the amount of bride wealth paid (and especially on the amount of *danga*, i.e. the most substation payment related to the rights over children, see Section 2). This variable needs to be interpreted with care. The information on bride wealth paid is affected by item non-response, and bride wealth paid is only observed at the household level, as the total amount of bride wealth obtained in the previous year. It is thus not related to a particular marriage. By combining the individual data on marriages with the household data on bride wealth, we relate bride wealth to marriages. In particular, if we observe the household receiving bride wealth we link this to the most recent marriage that occurred in this household. For the year 2000 we did not obtain any information on bride wealth.

On the basis of the bride wealth data, we create two variables. The first variable contains the amount of bride wealth received by the household in the year of marriage, which we refer to as short-term bride wealth. The second variable, which is called long-term bride wealth is the accumulated amount of bride wealth received by the household in the year of marriage and the following years in the observation period. However, recall that it usually takes around 30 years until the total amount of bride wealth has been paid, while we only observe yearly payments over a maximum of 5 years. The total amount of bride wealth we observe is therefore an underestimate of the actual total amount. We restrict our attention to households for which we observe household characteristics. For these households we observe 424 marriages in the data. However, we only find bride wealth information for 128 of them. In the sample the average bride wealth consisted of around 1300 Zimbabwean dollars (either in cash or in cattle).

With respect to long-term bride wealth we only include marriages until 1997, otherwise the total observed bride wealth would be too much affected by the remaining length of the observation period after the year of marriage. This subsample includes 174 marriages. For 156 of these 174 we observe bride wealth. The average bride wealth was approximately 2200 Zimbabwean dollar.

5 Estimation results

In this section we present the estimation results of the empirical analyses. We first discuss the estimation results obtained with the stock sample of unmarried women. After that we consider the results based on the flow sample of women reaching age 15 and compare both estimation results. Finally, estimation results concerning the amount of bride wealth are discussed.

5.1 Estimation results based on the stock sample

For estimating the parameters of the marriage rate using the stock sample, it is necessary to have information on the rate at which women reached age 15 prior to the observation period, $n(\tau)$ (see Subsection 3.1). Unfortunately, we do not observe any measure of inflow, or alternatively, the birth rate 15 years prior to τ. We therefore assume this to be constant, $n(\tau) = 1$, for each τ. We specify the piecewise constant baseline hazard in terms of two years and $\psi(\tau) = \exp(\delta r_{\tau,k})$ with $r_{\tau,k}$ denoting the quantity of rainfall in region k in the agricultural season before year τ. Hence, we estimate the parameters v_i, p_i $(i = 1, 2, \ldots, m)$, $\lambda_{j:j+1}$ $(j = 15, 17, \ldots, 29)$, β, which is a vector of parameters capturing the covariate effects, and δ, which is the effect of rainfall on the marriage rate. We normalize by taking $\lambda_{15:16} = 0$ and $p_m = 1 - p_1 - \ldots - p_{m-1}$. The model parameters are estimated using Maximum Likelihood.

Table 7 presents the parameter estimates obtained with the stock sample. We find two points of support of the distribution of the unobserved heterogeneity term $(m = 2)$. 78% of the probability mass is at one point. The remaining 22% of the women has a much lower marriage rate. The baseline hazard shows how the marriage rate is affected by the age of an unmarried woman. It shows that the marriage rate increases with age.

The main parameter of interest is δ, which represents the effect of rainfall on the marriage rate. The parameter estimate is negative as expected from the discussion in Section 2. It is not significantly different from 0. However, given the small sample size and the relatively small variation in local rainfall, it would be too ambitious to expect the convenient levels of significance. During the observation period the quantity of rainfall varied from around 0.5 meter in 1994 to around 1 meter in 1996 and 1998. This implies that in 1995 the marriage rate was on average only around 2.5% higher than in 1999, i.e. $\exp(0.05 \cdot 0.5) - 1 \approx 0.025$. The impact of rainfall on the marriage rate is thus very low. As mentioned before local rainfall is correlated with the average rainfall in Zimbabwe (e.g. Table 2). Like brides (and their families) grooms suffer from the negative weather shocks as well. After a drought they are likely to be reluctant to get married and to pay bride wealth. So even if households put pressure on their daughters to get married after a bad rainy season, the number of marriages might increase little because of the unavailability of grooms.

In the estimation we omitted the age and the years of education of the parents. The corresponding parameter estimates were very close to 0 and highly insignificant. The marriage rate is highest for women living in Mpfurundi. The women in the other regions on average marry later. Women with Masowe faith marry later than women with African faith or Christian faith.

Household characteristics may reveal to what extent the timing of marriage is a household decision. The size of the household has a positive effect on the marriage rate, implying that daughters living in larger households marry on average younger. This is intuitive, as for a large household the loss of labor due to a marriage is less severe. The amount of land is fixed, so the marginal productivity of a daughter in larger households is smaller than in smaller households. Furthermore, in case one of the parents is absent, the marriage rate is lower, particularly if the father is absent. We do not know the reason why a parent is absent, but recall that the father is much more often absent than the mother, in 17% and 4% of the households respectively. A father who is absent may work in the city, although as noted earlier finding work in the city is difficult. If the father is absent because of work in the city, the yearly income of the household is likely to be more stable. Therefore, household behavior is less sensitive to shocks and their need for cattle (as buffer stock) is reduced. The latter corresponds to the empirical fact that single-headed households are less well endowed with livestock than two-parent households. If the father works in the city, the household is less dependent on the income from agriculture, which means that the household does not need the labor of the children (although it might be that in single-headed households children have to perform (part of) the tasks of the missing parent). This affects the marginal productivity of a child in a single-headed family.

5.2 Estimation results based on the flow sample

Inference based on the stock sample does not allow us to investigate the effects of household wealth on the timing of the marriage. To investigate this we base inference on the flow sample. The specification of the hazard is to a large extent similar as in the previous subsection. However, the short observation period only allows us to estimate the baseline hazard until age 20. Furthermore, we now consider household size as a time-varying regressor and we include two additional regressors, the household's livestock wealth and a dummy variable for being the oldest unmarried daughter in the household. Since the flow sample does not contain a large number of women with absent mothers, we replace the dummy variables for an absent father and for an absent mother by a single dummy variable indicating that the household is single-headed.

The parameter estimates of the flow sample procedure are presented in Table 8. We do not find any significant unobserved heterogeneity. During the optimization of the loglikelihood function both mass points converge to a single mass point. In general, standard errors when using the flow sample are larger than when using the stock sample. This is not surprising as the flow sample only contains half of the women in the stock sample and has a much higher fraction right-censored spells. However, in most cases the parameters have the same sign as the stock sample parameter estimates. Again we see that the marriage rate increases with age and that women living in Mpfurudzi marry on average younger. Also faith and household size have about the same effects on the hazard rate as above. The significant effect of household size in the flow sample (where we also correct for household wealth) might indicate that for large families the consequences of the loss of labor following a girl leaving the household may not be as dramatic compared to small families. Alternatively, large households are likely to have many sons, which are (due to bride wealth) claims on household wealth. Ideally, a marriage of daughter precedes a marriage of a son, and thus women living in households with many brothers have more pressure to marry. Therefore, we have tried to replace household size with the number of sons in the household. This did not affect the parameter estimate. Both variables are too correlated to include them jointly.

Remarkable is that the effect of living in a single-headed household is positive which suggests that women in single-headed households marry earlier. As mentioned earlier single-headed households own less livestock than households with two parents; their livestock wealth is about 20% lower. The positive covariate effect of living in a single-headed household is therefore reinforced by the negative covariate effect of living in a poor household. This contradicts the estimation results obtained using the stock sample. As mentioned earlier, in single-headed households most often the father is absent. If the father is working in the city, the household is less dependent on agriculture and the labor of their children. The girls can just marry at a younger age.

The main difference between the estimation results of the stock sampling and the flow sampling is the effect of the amount of rainfall on the marriage rate. When using the flow sample, the effect of rainfall on the marriage rate is positive, implying that more marriages occur after a season with more rainfall. It is clear that in the flow sample model the effect of rainfall should be interpreted differently, as in the flow sample we also control for the household's livestock wealth. The amount of rainfall in this case is thus merely a marriage market effect, i.e. a season of low rainfall reduces the number of men available for marriage. The effect that due to a negative rainfall shock more women want to marry to compensate the reduction in household wealth, is now picked up by the livestock wealth variable. Finally, we should note that because of the flow sampling, most marriages occur in the last years of the observation period. Therefore, the estimated effect might be not very robust.

Now consider the two variables we added to the model. The effect of being the oldest unmarried daughter in the household is positive. The marriage rate of the oldest daughter is 57% $(= \exp(0.45) - 1)$ higher than the marriage of a girl of the same age, who has an older unmarried sister. This is in agreement with the hypothesis that the timing of marriage is a household decision rather than an individual decision. In case a household would like one of their daughters to get married, the oldest daughter is most likely to be pressed to do so. There are two main reasons. First, the amount of bride wealth is decreasing faster in age for older girls (if a women marries at young age, she will most likely have more children). So, the opportunity costs of staying unmarried for some period are highest for the oldest daughter in the household. Second, other daughters could easily argue against the parents' pressure to marry on the ground that they have an older unmarried sister. Since the model takes account of the age of women, this is not a spurious effect due to the fact that the average age of oldest daughters is most likely to be higher than of the other unmarried women.

The other additional regressor is livestock wealth, which has a negative effect on the marriage rate. Women from less wealthy households marry at a younger age than women from richer households. This is in accordance with the theory presented in Section 2, which argued that marriage of a daughter is a possibility to access livestock assets. To get some feeling for the magnitude of the effect of livestock wealth on the marriage rate, we compare the marriage rates in three households. The first household is a poor household that only has 2 cows, which might be just enough for plowing (if both are healthy). The second household has 5 cows, which guarantees enough draft power for this year, but is not considered as buffer stock. The third household is wealthy and has 8 cows, it thus has some buffer stock. The marriage rate of daughters in the first household is around 8.8% higher than the marriage rate in the second household and 18.3% higher than the marriage rate in the third household.

Based on these estimation results, it cannot be ruled out that this effect is due to the fact that daughters living in poor household have a strong incentive to leave the household, just to escape bad living conditions. To get some more insight in what drives the effect of livestock wealth on marriage, we extend the model by separating the livestock wealth in a fixed household component η, which is constant over time, and time-varying shocks ε_t in household wealth. We cannot simply take η equal to the mean household's livestock wealth over the years and ε_t as the difference between the household's livestock wealth in year t and the mean household's livestock wealth. If a marriage occurs the bride wealth payments increase the mean household's livestock wealth, which implies that before the moment of the marriage ε_t would be too low and after the moment of marriage too high. This might cause spurious relations between shocks and the marriage rate. Therefore, to obtain estimates for η and ε_t we estimate a dynamic panel data model for livestock wealth accumulation (that also takes marriages into account), η is the household specific effect in the model and ε are the disturbances in the model. The procedure is described in detail in Appendix A. The estimation results are given in Table 9. As can be seen, the effect of a shock in the livestock wealth is more important than differences between household specific components. This implies that the effect of livestock wealth on the marriage rate described above is induced by (unanticipated) shocks in livestock wealth rather than by existing wealth differences between households.

The effect of an (unanticipated) shock in the household's livestock wealth on the marriage rate is significant. If a household loses 2 head of cattle, due to an unanticipated event such as theft, the marriage rate of the daughters in this household increases with around 22%.

We have tried adding other explanatory variables to the model. However, this did not improve the fit, i.e. the parameter estimates were very close to 0 with relatively large standard errors and no improvement in the loglikelihood function. In particular, we added equipment owned by the household and total acres of land used in agricultural production. It should be stressed that these variables are highly correlated with livestock wealth, which implies that potential covariate effects of such regressors are already covered by livestock wealth.

5.3 Analyzing the amount of bride wealth payments

Finally, we performed some empirical analyses on the amount of the bride wealth. As mentioned in Section 4 the data on the amount of the bride wealth are not very precise and do not cover a sufficiently long period to observe the payment of total bride wealth as payments most often take as much as 30 years. The estimation results should therefore be interpreted with care. For the empirical analyses we use tobit models.[8] We perform two analyses, the first one based on the amount of bride wealth received during the year of marriage and the second based on the accumulated amount of bride wealth received during the observation period after the marriage. The latter also includes the bride wealth received in the year of marriage. As explanatory variables we include age of the daughter

[8]Since we only observe the amount of bride wealth for a subsample of the data, the ideal model would be a sample selection model. However, identification of sample selection models hinges on exclusion restrictions. In our data there is no variable that would qualify for being excluded from the equation denoting the amount of bride wealth, but included in the equation describing if the household reports having received bride wealth. Therefore, we use a simpler censored regression or tobit model,

$$y^* = x\beta + \varepsilon$$

and

$$y = \begin{cases} y^* & \text{if} \quad y^* > 0 \\ 0 & \text{if} \quad y^* \leq 0 \end{cases}$$

The interpretation of this model is that bride wealth is observed with a large measurement error ε, and that households only report having received bride wealth if the bride wealth including the measurement error exceeds 0.

at the moment of marriage, amount of rainfall, an indicator for being the oldest daughter and livestock wealth in the year before marriage.

The estimation results are given in Table 10. Except for rainfall all covariate effects on the amount of bride wealth received in the year of marriage, are opposite to the effect on the total amount of the bride wealth. Obviously, some households prefer a relatively large amount of bride wealth at the moment of marriage over a larger over-all amount of bride wealth. The amount of bride wealth received at the first year is higher in relatively dry years, for poorer households and if it is the oldest daughter who marries. Above we already explained that since we also correct for age, the dummy variable for oldest daughter can be interpreted as an indicator for household pressure. It is clear that households who bargain for a relatively large initial payment of bride wealth, have to accept low payments in the following years. This observation is in accordance with the discussion in Section 2, that poor households can escape a period of poverty by the marriage of a daughter, but suggests that after the substantial initial transfer a period with low transfers follows.

6 Conclusions

This paper focuses on the timing of marriage in rural areas in Zimbabwe. In these rural areas farmers live who obtained 12 acres of land from a resettlement program in the early 1980s. To work the land cattle is required; without cattle agricultural production is minimal. The absence of access to formal financial institutions to smallholder farmers causes that households cannot borrow to buy cattle if they do not possess sufficient cattle for plowing. Since finding off-farm jobs is extremely difficult and the return to manually working the fields is low, households without cattle may get stuck in poverty for a long period. Due to the existence of bride wealth, which is a transfer from (the family of) the groom to the family of the bride, a daughter's marriage is a possibility to acquire cattle. An unmarried daughter could therefore be considered as an asset that can be cashed in bad times.

In the empirical analyses we focus on the influence of weather shocks and (shocks to) household cattle wealth on the timing of a daughter's marriage. If indeed the marriage behavior depends on shocks, the composition of the population of unmarried women in a year depends on shocks in the previous years. To take account of these dynamics, we use a mixed proportional hazard rate framework to model the age of marriage.

By focusing on the timing of a marriage of a daughter, we only provide a partial picture on the marriage market. We do not observe any information on the choice of spouses. However, the lack of information on the choice of spouses makes it impossible to extend this picture.

For the empirical analyses we rely on a unique panel data set on Zimbabwean smallholder farmers. Due to all kinds of shocks the wealth of the smallholder farmers fluctuates enormously over the years, i.e. being wealthy in a particular does not guarantee being wealthy in next years.

The estimation results show that in the stock sampling model the impact of rainfall on the rate at which women marry is negligible. However, if we also control for the household livestock wealth, the amount of rainfall has a positive effect on the marriage rate. This is most likely a marriage market effect, i.e. after a period of low rainfall the supply of men is lower as not that many households are capable of paying bride wealth. The marriage rate of daughters is higher for poorer households. In particular, after a negative shock in livestock wealth the marriage rate of daughters increases. Furthermore, the marriage rate for the oldest daughter in the household is higher. Both these effects confirm the hypothesis that households use an unmarried daughter as an asset that can be cashed in bad times. The marriage decision should therefore not only be considered as a decision of the individuals involved, but also as a household decision.

We also considered the amount of bride wealth. These estimation results provide additional evidence that the timing of marriage is used as means to avoid being stuck in poverty. In particular, less wealthy households receive a larger initial payment at the moment a daughter gets married. The accumulated amount of bride wealth they receive in the first years of marriage is relatively low, so that the relatively high initial payment goes at the expense of subsequent payments. In other words, a period of poverty is only avoided at the expenses of receiving lower payments in the subsequent years.

Our results have policy relevance. In particular, the use of marriage as an alternative financial institution affects the age at which daughters marry. This might have negative external effects. If women marry young, they are more likely to have more children. Furthermore, households might want to have at least a particular number of daughters to act as insurance against negative shocks. This induces households to have more children (especially if the first children born to the family are boys).

References

Binswanger, H.P. and J. McIntire (1987), Behavioural and material determinants of production relations in land-abundant tropical agriculture, *Economic Development and Cultural Change* 36, 73–99.

Binswanger, H.P. and M.R. Rosenzweig (1986), Behavioural and material determinants of production relations in agriculture, *Journal of Development Studies* 22, 503–539.

Carter, M.R. and F.J. Zimmerman (2000), The dynamic cost and persistence of asset inequality in an agrarian economy, *Journal of Development Economics* 63, 265–302.

Central Statistical Office (1995), *Demographic and Health Survey 1994*, Harare, Zimbabwe.

Deininger, K. and J.G.M Hoogeveen (2002) Economic Costs and Benefits of Land Redistribution in Zimbabwe *World Development*, forthcoming.

Dekker, M. and J.G.M. Hoogeveen (2002), Bride wealth and household security in rural zimbabwe, *Journal of African Economies* 11, 114–145.

Fafchamps, M., C. Udry and K. Czukas (1998), Drought and saving in West-Africa: Are livestock a buffer stock?, *Journal of Development Economics* 55, 273–305.

Flinn, C. and J.J. Heckman (1982), New Methods for Analyzing Structural Models of Labor Force Dynamics, *Journal of Econometrics* 18, 115–168.

Jacoby, H.G. (1995), The economics of polygyny in sub-Saharan Africa: female productivity and the demand for wives in Côte d'Ivoire, *Journal of Political Economy* 103, 938–971.

Kinsey, B., K. Burger and J.W. Gunning (1998), Coping with drought in Zimbabwe: Survey evidence on responses of rural households to risk, *World Development* 1, 89–110

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.

Rao, V. (1993), The rising price of husbands: A hedonic analysis of dowry increases in rural India, *Journal of Political Economy* 101, 666–677.

Ridder, G. (1984), The distribution of single-spell duration data, in G.R. Neumann and N. Westergård-Nielsen (eds.), *Studies in Labor Market Analysis*, Springer-Verlag, Berlin.

Rosenzweig, M.R. (1988), Risk, implicit contracts and the family in rural areas of low income countries, *Economic Journal* 98, 1148–1170.

Rosenzweig, M.R. (2001), Savings behaviour in low-income countries, *Oxford Review of Economic Policy* 17, 40–54.

Rosenzweig, M.R. and O. Stark (1989), Consumption smoothing, migration, and marriage: evidence from rural India, *Journal of Political Economy* 97, 905–926.

Rosenzweig, M.R. and K.I. Wolpin (1993), Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: investment in bullocks in India, *Journal of Political Economy* 101, 223–244.

Scoones, I. (1995), Investigating difference: Applications of wealth ranking and household survey approaches among farming households in southern Zimbabwe, *Development and Change* 26, 67–88.

Scoones, I., C. Chibudu, S. Chikura, P. Jeranyama, D. Machaka, W. Machanja, B. Mavedzenge, B. Mombeshora, M. Mudhara, C. Mudziwo, F. Murimbarimba and B. Zirereza (1996), *Hazards and Opportunities, Farming Livelihoods in Dryland Africa: Lessons from Zimbabwe*, Zed Books Ltd., London and New Jersey.

Stern, S. (1997), Simulation-based estimation, *Journal of Economic Literature* 35, 2006-2039.

Tsodzo, T.K.M. (1992), Cattle are our bank, in C. Kileff and P. Kileff (eds.), *Shona customs, essays by African writers*, Mambo Press, Gweru, Zimbabwe.

A Separating household specific effects and shocks

In this appendix we briefly discuss our method for estimating the marriage rate that includes a fixed household component for livestock wealth and a component that describes idiosyncratic shocks in the household's livestock wealth. We obtain information on both components from a dynamic panel data model for livestock wealth accumulation. First, we provide and estimate the model for livestock wealth accumulation and next we discuss how these estimation results are used in estimating the marriage rate.

A.1 Dynamic model for livestock wealth accumulation

Our data set contains information on H households that are denoted by $h = 1, \dots, H$. The panel is unbalanced, however, the number of years T_h for which we observe household h is exogenous. For a particular year $t = 1, \dots, T_h$, the household's livestock wealth $w_{h,t}$ (measured in cows) depends on (i) the household's livestock wealth one year earlier $w_{h,t-1}$, (ii) a dummy variable $m_{h,t}$ that indicates if a daughter in the household married between the survey in year $t-1$ and year t,[9] (iii) the size of the household $s_{h,t-1}$ during the survey in year $t-1$, (iv) the amount of rainfall r_{t-1} in year $t-1$, and (v) some household characteristics z_h that are constant over time. These latter household characteristics include dummy variables for living in the Mutanda and the Sengezi region, for having the African faith and Masowe faith, and for living in a single-headed household.

The household's livestock wealth accumulates according to the model

$$w_{h,t} = \beta_0 + \beta_1 w_{h,t-1} + \beta_2 m_{h,t} + \beta_3 s_{h,t-1} + \beta_4 r_{t-1} + z_h \beta_5 + \eta_h + \varepsilon_{h,t},$$

where η_h is an unobserved household specific wealth component and $\varepsilon_{h,t}$ are random shocks. These are the two components that we want to include as regressors in the marriage rate. We assume that both η_h and $\varepsilon_{h,t}$ follow a normal distribution with mean zero and variance σ_η^2 and σ_ε^2 respectively.

The household specific wealth component η_h is not independent of $w_{h,t-1}$. To

[9] As mentioned earlier not all bride wealth is paid in the year of marriage, which might suggest that we should also include the dummy variables for marriages lagged. However, bride wealth payments are most substantial in the year of marriage and bride wealth payment schemes are not similar. Furthermore, we do not observe marriages occurring before 1994, so including these lagged variable would reduce the length of our panel considerably.

get rid of η_h, we take first differences,

$$\Delta w_{h,t} = \beta_1 \Delta w_{h,t-1} + \beta_2 \Delta m_{h,t} + \beta_3 \Delta s_{h,t-1} + \beta_4 \Delta r_{t-1} + \Delta \varepsilon_{h,t}$$

However, the new regressor $\Delta w_{h,t-1}$ obtained after the transformation is correlated with the error term $\Delta\varepsilon_{h,t}$ (as $w_{h,t-1}$ depends on $\varepsilon_{h,t-1}$). Also marriage $m_{h,t}$ might depend on shocks in household's livestock wealth one year earlier, which suggests that $m_{h,t}$ can be correlated with $\varepsilon_{h,t-1}$. Therefore, we will use two-stage least squares (2SLS) to estimate the model, where in the first stage we estimate

$$\begin{aligned}
\Delta w_{h,t-1} = {} & \gamma_0 + \gamma_1 w_{h,t-2} + \gamma_2 w_{h,t-3} + \gamma_3 m_{h,t-1} + \gamma_4 \Delta s_{h,t-1} + \gamma_5 \Delta s_{h,t-2} \\
& + \gamma_6 \Delta r_{t-1} + \gamma_7 \Delta r_{t-2} + u^1_{h,t} \\
\Delta m_{h,t} = {} & \delta_0 + \delta_1 w_{h,t-2} + \delta_2 w_{h,t-3} + \delta_3 m_{h,t-1} + \delta_4 \Delta s_{h,t-1} + \delta_5 \Delta s_{h,t-2} \\
& + \delta_6 \Delta r_{t-1} + \delta_7 \Delta r_{t-2} + u^2_{h,t}
\end{aligned}$$

This 2SLS procedure provides consistent estimators $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ (as $H \to \infty$).[10] As an estimator for σ^2_ε, we use

$$\hat{\sigma}^2_\varepsilon = \frac{1}{H} \sum_{h=1}^{H} \frac{1}{T_h - 1} \sum_{t=1}^{T_h} \left(\hat{u}_{h,t} - \frac{1}{T_h} \sum_{t=1}^{T_h} \hat{u}_{h,t} \right)^2$$

where

$$\hat{u}_{h,t} = w_{h,t} - \hat{\beta}_1 w_{h,t-1} - \hat{\beta}_2 m_{h,t} - \hat{\beta}_3 s_{h,t-1} - \hat{\beta}_4 r_{t-1}$$

When computing the standard errors for $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\beta}_4$ we have corrected for correlation between $\Delta\varepsilon_{h,t}$ and $\Delta\varepsilon_{h,t-1}$, i.e. $cov(\Delta\varepsilon_{h,t-1}, \Delta\varepsilon_{h,t}) = -\sigma^2_\varepsilon$ and $var(\Delta\varepsilon_{h,t}) = 2\sigma^2_\varepsilon$. The estimation results are given in Table 11.

Next, we estimate the parameters β_0 and β_5. Therefore, we use the regression

$$\frac{1}{T_h} \sum_{t=1}^{T_h} \hat{u}_{h,t} = \beta_0 + z_h \beta_5 + e_h$$

where it can be shown that

$$\begin{aligned}
e_h = {} & \eta_h + \frac{1}{T_h} \sum_{t=1}^{T_h} \varepsilon_{h,t} + \frac{1}{T_h} \sum_{t=1}^{T_h} (\beta_1 - \hat{\beta}_1) w_{h,t-1} + (\beta_2 - \hat{\beta}_2) m_{h,t} \\
& + (\beta_3 - \hat{\beta}_3) s_{h,t-1} + (\beta_4 - \hat{\beta}_4) r_{t-1}
\end{aligned}$$

[10] We have performed a Sargan-test to test the specification of our model and the validity of the set of instrumental variables. The value of the test statistics equals 4.51. Since it follows a χ^2-distribution with 3 degrees of freedom, we cannot reject that the model is misspecified or that the instrumental variables are invalid.

We can use OLS to estimate the parameters β_0 and β_5, but when computing standard errors we should take into account that the disturbances e_h are all correlated with each other and that they suffer from heteroskedasticity. The estimation results are reported in Table 12.

As will become clear below, we need to know the complete variance-covariance matrix of $\left[\hat{\beta}_0; \hat{\beta}_5'\right]'$ and $\left[\hat{\beta}_1; \hat{\beta}_2; \hat{\beta}_3; \hat{\beta}_4\right]'$. This implies that we also should estimate the covariance matrix of the estimators $\left[\hat{\beta}_0; \hat{\beta}_5'\right]'$ with the estimators $\left[\hat{\beta}_1; \hat{\beta}_2; \hat{\beta}_3; \hat{\beta}_4\right]'$. We know that

$$\begin{bmatrix} \hat{\beta}_0 - \beta_0 \\ \hat{\beta}_5 - \beta_5 \end{bmatrix} = \left([\iota; Z]' [\iota; Z]\right)^{-1} [\iota; Z]' \begin{bmatrix} e_1 \\ \vdots \\ e_H \end{bmatrix}$$

where ι is a vector containing 1s and Z is a matrix with the vectors z_h. As we know how e_h depends on $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, and $\hat{\beta}_4$, we get an expression for the covariance matrix between $\left[\hat{\beta}_0; \hat{\beta}_5'\right]'$ and $\left[\hat{\beta}_1; \hat{\beta}_2; \hat{\beta}_3; \hat{\beta}_4\right]'$. Let the estimated variance-covariance matrix of the estimators $\left[\hat{\beta}_0; \hat{\beta}_1; \hat{\beta}_2; \hat{\beta}_3; \hat{\beta}_4'; \hat{\beta}_5\right]'$ be given by $\hat{\Omega}$.

A.2 Including a household specific effect and shocks as regressors in the marriage rate.

The marriage rate is specified at the level of daughters instead of households. Woman $i = 1, \ldots, N$ lives in a household with household specific effect η_i and yearly shocks are $\varepsilon_{i,1}, \ldots, \varepsilon_{i,T}$. The characteristics of the women are given by $x_{i,\tau}$, where τ denotes the year in which these are observed. As shown in Section 3.2 the likelihood function for the flow sample equals

$$\prod_{i=1}^{N} \Pr(T \in [\tau_i^*, \tau_i^* + 1) | \text{ Inflow at } \tau_{s,i}, x_{i,\tau_s}, \ldots, x_{i,\tau^*}, \eta_i, \varepsilon_{i,\tau_s}, \ldots, \varepsilon_{i,\tau^*})$$

If we would know the values of η_i and $\varepsilon_{i,\tau}$, we could just optimize the logarithm of this likelihood function to obtain the parameter estimates. However, instead of η_i and $\varepsilon_{i,\tau}$ we observe the estimated residuals

$$\hat{v}_{h,t} = w_{h,t} - \hat{\beta}_0 - \hat{\beta}_1 w_{h,t-1} - \hat{\beta}_2 m_{h,t} - \hat{\beta}_3 x_{h,t-1} - \hat{\beta}_4 r_{t-1} - z_h \hat{\beta}_5$$

We should therefore optimize the loglikelihood function

$$\log\left(\int \prod_{i=1}^{N} \Pr(T \in [\tau_i^*, \tau_i^* + 1) | \text{ Inflow at } \tau_{s,i}, x_{i,\tau_s}, \ldots, x_{i,\tau^*}, \eta_i, \varepsilon_{i,\tau_s}, \ldots, \varepsilon_{i,\tau^*}) \right.$$
$$\left. f(\eta, \varepsilon|\hat{v}) \ \mathrm{d}(\eta, \varepsilon)\right)$$

where η is a vector containing all household specific effects, ε is a matrix containing for all households all yearly shocks and $\hat{v}$ is a matrix containing the elements $\hat{v}_{h,t}$.

The remaining problem is to specify the density function $f(\eta, \varepsilon|\hat{v})$. We rewrite this density function by conditioning on a matrix v containing the elements $v_{h,t} = \eta_h + \varepsilon_{h,t}$, as

$$f(\eta, \varepsilon|\hat{v}) = \int f(\eta, \varepsilon|v, \hat{v}) f(v|\hat{v}) dv$$

It is clear that if we know v, then $\hat{v}$ is not informative about η and ε, thus

$$f(\eta, \varepsilon|v, \hat{v}) = f(\eta, \varepsilon|v)$$

This density function can be written as

$$\begin{aligned} f(\eta, \varepsilon|v) &= \prod_{h=1}^{H} f(\eta_h, \varepsilon_{h,1} \ldots, \varepsilon_{h,T_h}|v_{h,1}, \ldots, v_{h,T_h}) \\ &\propto \prod_{h=1}^{H} f(v_{h,1}, \ldots, v_{h,T_h}|\eta_h, \varepsilon_{h,1}, \ldots, \varepsilon_{h,T_h}) f(\eta_h, \varepsilon_{h,1}, \ldots, \varepsilon_{h,T_h}) \\ &= \prod_{h=1}^{H} \phi(\eta_h) \prod_{t=1}^{T_h} I(\varepsilon_{h,t} = v_{h,t} - \eta_h)\phi(\varepsilon_{h,t}) \end{aligned}$$

where $I(\cdot)$ is the indicator function. In the last step we used that both η_i and $\varepsilon_{i,t}$ follow normal distributions (as was assumed earlier).

Next we have to focus on $f(v|\hat{v})$. Note that

$$\begin{aligned} v_{h,t} =& \ w_{h,t} - \beta_0 - \beta_1 w_{h,t-1} - \beta_2 m_{h,t} - \beta_3 s_{h,t-1} - \beta_4 r_{t-1} - z_h\beta_5 \\ =& \ \hat{\beta}_0 + \hat{\beta}_1 w_{h,t-1} + \hat{\beta}_2 m_{h,t} + \hat{\beta}_3 s_{h,t-1} + \hat{\beta}_4 r_{t-1} + z_h\hat{\beta}_5 + \hat{v}_{h,t} \\ & - \beta_0 - \beta_1 w_{h,t-1} - \beta_2 m_{h,t} - \beta_3 s_{h,t-1} - \beta_4 r_{t-1} - z_h\beta_5 \\ =& \ (\hat{\beta}_0 - \beta_0) + (\hat{\beta}_1 - \beta_1) w_{h,t-1} + (\hat{\beta}_2 - \beta_2) m_{h,t} \\ & + (\hat{\beta}_3 - \beta_3) s_{h,t-1} + r_{t-1}(\hat{\beta}_4 - \beta_4) + z_h(\hat{\beta}_5 - \beta_5) + \hat{v}_{h,t} \end{aligned}$$

This implies that the vectorized matrix v is asymptotically normal distributed with mean the vectorized matrix $\hat{v}$ and variance-covariance matrix $X\hat{\Omega}X'$, where X is a matrix that contains rows $[1; w_{h,t-1}; m_{h,t}; s_{h,t-1}; r_{t-1}; z_h]$.

To optimize the loglikelihood function given above we use simulated maximum likelihood estimation (e.g. Stern, 1997). We draw 1000 times from the distribution $f(\eta, \varepsilon|\hat{v})$. For each draw $j = 1, \ldots, 1000$ we follow procedure: (i) given the values $\hat{v}$ we generate a matrix v_j, (ii) we generate a vector η_j with household specific effects from a normal distribution with mean 0 and variance $\hat{\sigma}_\eta^2$, (iii) we compute the matrix of shocks $\varepsilon_j = v_j - \eta_j$, (iv) we compute the likelihood contribution and we weight this by $p_j = \phi(\varepsilon_j)$, (v) we sum the weighted individual likelihood contributions and optimize its logarithm. The estimation results are given in Table 9

Cattle type	Survival (%) 1987/1988	Survival (%) 1991/1992
Bulls	90.6	34.0
Oxen	96.0	33.6
Cows	71.3	26.9
Heifers	76.3	76.9
Male calf	81.4	37.5
Female calf	58.3	41.2

Source: Scoones, Chibudu, Chikura, Jeranyama, Machaka, Machanja, Mavedzenge, Mombeshora, Mudhara, Mudziwo, Murimbarimba and Zirereza (1996).

Table 1: Cattle survival rates for beasts kept near Chivi in South Zimbabwe.

Year	Sample size at beginning of year	Number of marriages	Average age of marriage	Rainfall in previous season local	Rainfall in previous season Zimbabwe
Mpfurudzi					
1994	215	34	20.5	0.69	0.52
1995	209	29	19.9	0.52	0.42
1996	202	26	19.3	0.77	0.70
1997	182	25	19.4	1.25	0.75
1998	180	17	18.4	0.70	0.53
1999	169	25	18.9	0.99	0.78
Mutanda					
1994	84	9	21.1	0.72	0.52
1995	80	7	23.3	0.58	0.42
1996	86	10	18.4	0.93	0.70
1997	76	11	19.0	0.97	0.75
1998	63	3	18.3	0.70	0.53
1999	58	3	19.0	1.27	0.78
Sengezi					
1994	60	7	20.0	0.52	0.52
1995	60	9	18.1	0.47	0.42
1996	55	9	19.9	0.90	0.70
1997	51	5	20.8	0.96	0.75
1998	44	1	16.0	0.72	0.53
1999	39	3	17.7	0.95	0.78
Total		233			

Explanatory note: The quantity of rainfall is measured in meters per year and collected by the Department of Meteorological Services in Zimbabwe.

Table 2: Some annual statistics of the sample of unmarried women in the age interval between 15 and 30.

	Head of cattle		
	0-2	3-4	5+
1994	20.2	21.4	22.0
1995	19.1	22.4	21.4
1996	20.7	20.8	19.8
1997	20.4	19.3	21.0
1998	18.9	19.5	19.8
1999	19.2	17.5	19.7

Table 3: Average age at which a daughter gets married stratified by the year of marriage and the herd size (cows, bulls, trained oxen and heifers) of the household.

Variable	Average
Region	
Mpfurudzi (%)	63
Mutanda (%)	21
Sengezi (%)	16
Individual characteristics	
African faith (%)	17
Masowe faith (%)	6
Household characteristics	
Missing (%)	10
Household size (persons)	12
Father absent (%)	17
Mother absent (%)	4
Age father (years)	53
Age mother (years)	45
Years of education father	5.1
Years of eduction mother	4.1
Number of observations	691

Explanatory note: Age of the father and the mother are measured at the moment the daughter was 15 years of age. Household size is measured at the moment the women first appeared in the stock sample.

Table 4: Stock sample characteristics.

Year	Sample size at beginning of year	Number of marriages	Average age of marriage
1994	62	0	–
1995	122	1	16.0
1996	175	8	16.3
1997	204	13	17.1
1998	219	12	17.3
1999	221	23	18.6
Total		57	

Table 5: Some annual statistics of the flow sample of unmarried women.

Year	Livestock wealth	
	Not married	Married
1994	9.7	–
1995	11.3	0.5
1996	12.0	6.6
1997	12.7	10.7
1998	13.6	9.1
1999	13.4	14.7

Explanatory note: In 1995, the prices (in Zimbabwean dollars) of cattle were: Cows 1200; Heifer 1000; Trained ox 1800; Young ox 1000; Bull 1500; and Goat 85.

Table 6: Average livestock wealth (in its real value in 1995) owned by the household, stratified by the women getting married in a particular year or not.

	Marriage hazard	
Unobserved heterogeneity		
v_1	−2.41	(0.41)
v_2	−4.15	(0.83)
p_1	0.78	(0.41)
p_2	0.22	(0.11)
Baseline hazard (Age)		
$\lambda_{15:16}$	0	
$\lambda_{17:18}$	0.70	(0.33)
$\lambda_{19:20}$	1.23	(0.30)
$\lambda_{21:22}$	1.79	(0.36)
$\lambda_{23:24}$	1.98	(0.61)
λ_{25+}	2.90	(0.82)
Rainfall (in meters)		
	−0.050	(0.18)
Region		
Mpfurundi	0	
Mutanda	−0.42	(0.19)
Sengezi	−0.21	(0.19)
Individual characteristics		
African faith	−0.094	(0.22)
Masowe faith	−0.20	(0.27)
Household characteristics		
Missing	0.44	(0.29)
Household size	0.021	(0.014)
Father absent	−0.27	(0.23)
Mother absent	−0.17	(0.30)
$\log \mathcal{L}$	-2257.49	
N	691	

Explanatory note: Standard errors in parentheses.

Table 7: Estimation results of the model using the stock sampling scheme.

	Marriage hazard	
Intercept		
v	-5.00	(0.94)
Baseline hazard (Age)		
λ_{15-16}	0	
λ_{17-18}	1.05	(0.39)
λ_{19-20}	2.11	(0.47)
Rainfall (in meters)		
	0.77	(0.77)
Region		
Mpfurudzi	0	
Mutanda	-0.31	(0.38)
Sengezi	-0.46	(0.43)
Individual characteristics		
African faith	-0.61	(0.63)
Masowe faith	-0.35	(0.58)
Oldest daughter	0.45	(0.33)
Household characteristics		
Missing	1.48	(0.56)
Household size	0.080	(0.025)
Parent absent	0.38	(0.38)
Livestock wealth (in cows)	-0.028	(0.019)
$\log \mathcal{L}$	-186.90	
N	333	

Explanatory note: Standard errors in parentheses.

Table 8: Estimation results of the model using the flow sampling scheme.

	Marriage hazard	
Intercept		
v	−5.38	(0.99)
Baseline hazard (Age)		
λ_{15-16}	0	
λ_{17-18}	1.04	(0.40)
λ_{19-20}	2.02	(0.52)
Rainfall (in meters)		
	0.64	(0.77)
Region		
Mpfurudzi	0	
Mutanda	−0.26	(0.39)
Sengezi	−0.50	(0.45)
Individual characteristics		
African faith	−0.78	(0.67)
Masowe faith	−0.40	(0.61)
Oldest daughter	0.51	(0.35)
Household characteristics		
Missing	1.97	(0.56)
Household size	0.083	(0.027)
Parent absent	0.41	(0.39)
Livestock wealth shock ($\varepsilon_{i,t}$ in cows)	−0.10	(0.049)
Livestock wealth level (η_i in cows)	−0.025	(0.026)
$\log \mathcal{L}$	-185.76	
N	333	

Explanatory note: Standard errors in parentheses.

Table 9: Estimation results using the flow sample of the model that distinguishes between household specific livestock wealth effects and (unanticipated) shocks in household's livestock wealth.

	Bride wealth payments			
	Short term		Long term	
Intercept	1.02	(3.10)	1.73	(1.35)
Age	−1.14	(0.090)	0.024	(0.051)
Rainfall	−1.85	(3.11)	−1.21	(1.51)
Oldest daughter	0.72	(0.62)	−0.14	(0.41)
Livestock wealth (in cows)	−0.043	(0.035)	0.027	(0.019)
σ	0.93	(0.037)	2.08	(0.097)
N	424		174	

Explanatory note: Standard errors in parentheses.

Table 10: Estimation results of the censored regression (tobit) model for the amount of bride wealth (in cows) received in the year of marriage (short term) and received in total (long term).

	$\Delta w_{h,t}$			
$\Delta w_{h,t-1}$	β_1	0.053	(0.086)	
$\Delta m_{h,t}$	β_2	0.81	(0.64)	
$\Delta s_{h,t-1}$	β_3	−0.0061	(0.098)	
$\Delta r_{h,t-1}$	β_4	0.73	(0.64)	
	σ_ε^2	21.4		

	$\Delta w_{h,t-1}$			$\Delta m_{h,t}$		
Intercept	γ_0	0.80	(0.28)	δ_0	0.12	(0.019)
$w_{h,t-2}$	γ_1	−0.43	(0.028)	δ_1	−0.00064	(0.0019)
$w_{h,t-3}$	γ_2	0.42	(0.031)	δ_2	0.00025	(0.0020)
$m_{h,t-1}$	γ_3	0.70	(0.48)	δ_3	−0.92	(0.032)
$\Delta s_{h,t-1}$	γ_4	0.091	(0.073)	δ_4	0.0086	(0.0048)
$\Delta s_{h,t-2}$	γ_5	0.14	(0.063)	δ_4	0.011	(0.0041)
$\Delta r_{h,t-1}$	γ_6	0.19	(0.55)	δ_4	−0.0032	(0.036)
$\Delta r_{h,t-2}$	γ_7	0.36	(0.57)	δ_4	0.047	(0.037)
R^2	0.20			0.48		
$F_{7,998}$	34.7			128.91		
H	290					
$\sum_{h=1}^{H}(T_h-1)$	1006					

Explanatory note: (corrected) standard errors in parentheses.

Table 11: Estimation results of 2SLS for β_1, β_2, β_3, and β_4.

Intercept	β_0	10.53	(2.78)
Mutunda	β_5	−2.09	(1.31)
Sengezi	β_5	0.61	(1.35)
Parent absent	β_5	−1.67	(1.25)
African faith	β_5	2.05	(1.46)
Masowe faith	β_5	−0.058	(2.19)
	σ_η^2	69.8	
H	290		

Explanatory note: (corrected) standard errors in parentheses.

Table 12: Estimation results of OLS for β_0 and β_5.

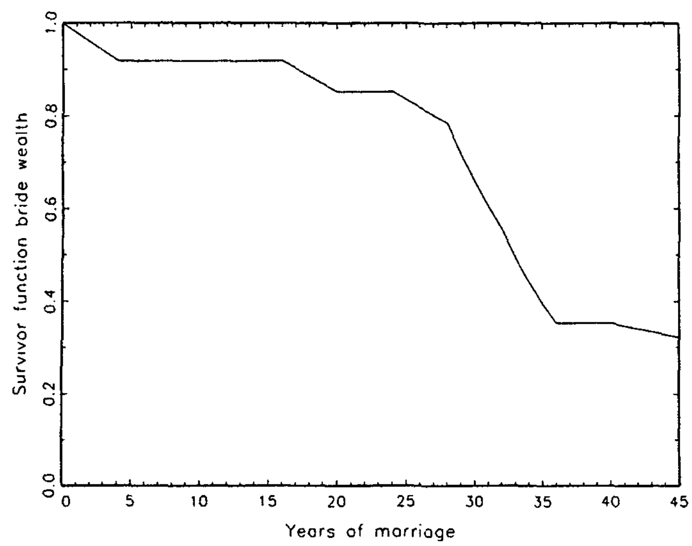

Figure 1: The fraction of the marriages where all bride wealth payments are made at a given number of years of marriage. This figure is based on a supplement held in 1995 of the data used in this paper.

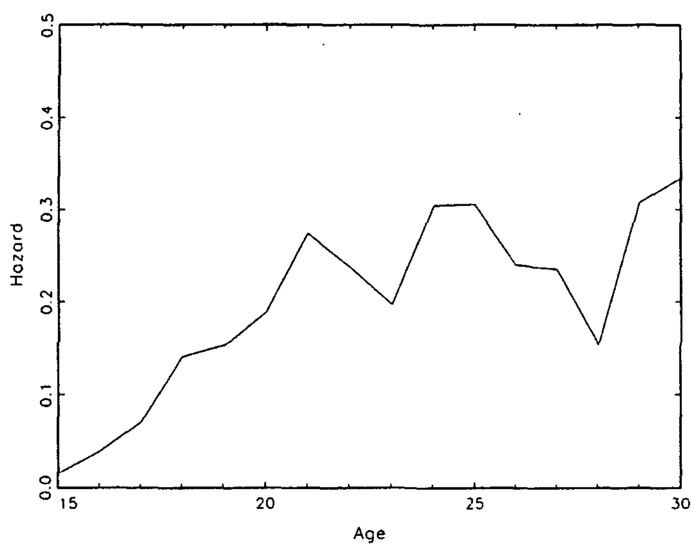

Figure 2: The empirical marriage rate, i.e. the probability of getting married at a given age conditional on still being unmarried when reaching this age.

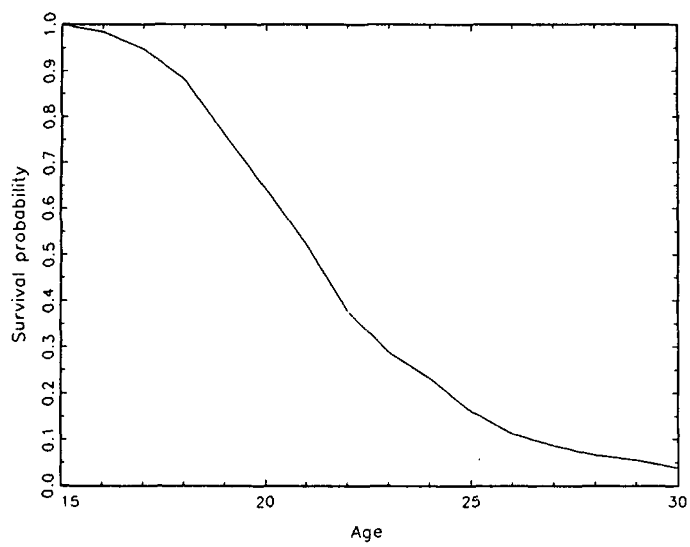

Figure 3: Kaplan-Meier estimate of the survivor function. This shows the percentage of women still unmarried at a given age.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS3088	U.S. Contingent Protection against Honey Imports: Development Aspects and the Doha Round	Julio J. Nogués	June 2003	P. Flewitt 32724
WPS3089	The "Glass of Milk" Subsidy Program and Malnutrition in Peru	David Stifel Harold Alderman	June 2003	H. Sladovich 37698
WPS3090	The Cotonou Agreement and its Implications for the Regional Trade Agenda in Eastern and Southern Africa	Manuel de la Rocha	June 2003	F. Sy 89750
WPS3091	Labor Market Policies and Unemployment in Morocco: A Quantitative Analysis	Pierre-Richard Agénor Karim El Aynaoui	July 2003	M. Gosiengfiao 33363
WPS3092	The Integrated Macroeconomic Model for Poverty Analysis: A Quantitative Macroeconomic Framework for the Analysis of Poverty Reduction Strategies	Pierre-Richard Agénor Alejandro Izquierdo Hippolyte Fofack	July 2003	M. Gosiengfiao 33363
WPS3093	Migration and Human Capital in Brazil during the 1990s	Norbert M. Fiess Dorte Verner	July 2003	R. Izquierdo 84161
WPS3094	Oil, Agriculture, and the Public Sector: Linking Intersector Dynamics in Ecuador	Norbert M. Fiess Dorte Verner	July 2003	R. Izquierdo 84161
WPS3095	The Emerging Project Bond Market: Covenant Provisions and Credit Spreads	Mansoor Dailami Robert Hauswald	July 2003	S. Crow 30763
WPS3096	International Survey of Integrated Financial Sector Supervision	José de Luna Martínez Thomas A. Rose	July 2003	S. Coca 37474
WPS3097	Measuring Up: New Directions for Environmental Programs at the World Bank	Piet Buys Susmita Dasgupta Craig Meisner Kiran Pandey David Wheeler Katharine Bolt Kirk Hamilton Limin Wang	July 2003	Y. D'Souza 31449
WPS3098	Governance and Economic Growth	Mark Gradstein	July 2003	H. Sladovich 37698
WPS3099	Creating Partnerships for Capacity Building in Developing Countries: The Experience of the World Bank	F. Desmond McCarthy William Bader Boris Pleskovic	July 2003	D. Philage 36971
WPS3100	Governance of Communicable Disease Control Services: A Case Study and Lessons from India	Monica Das Gupta Peyvand Khaleghian Rakesh Sarwal	July 2003	H. Sladovich 37698
WPS3101	Portfolio Preferences of Foreign Institutional Investors	Reena Aggarwal Leora Klapper Peter D. Wysocki	July 2003	A. Yaptenco 31823
WPS3102	Investing in Infrastructure: What is Needed from 2000 to 2010?	Marianne Fay Tito Yepes	July 2003	M. Fay 87200

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS3103	Ownership Structure and Initial Public Offerings	Reena Aggarwal Leora Klapper	July 2003	A. Yaptenco 31823
WPS3104	Does Strict Employment Protection Discourage Job Creation? Evidence from Croatia	Jan Rutkowski	August 2003	J. Rutkowski 84569
WPS3105	Further Evidence on the Link between Finance and Growth: An International Analysis of Community Banking and Economic Performance	Allen N. Berger Iftekhar Hasan Leora F. Klapper	August 2003	A, Yaptenco 31823
WPS3106	Governance Matters III: Governance Indicators for 1996–2002	Daniel Kaufmann Aart Kraay Massimo Mastruzzi	August 2003	A. Bonfield 31248
WPS3107	More Favorable and Differential Treatment of Developing Countries: Toward a New Approach in the World Trade Organization	Bernard Hoekman Constantine Michalopoulos L. Alan Winters	August 2003	P. Flewitt 32724
WPS3108	Stabilization and Association Process in the Balkans: Integration Options and their Assessment	Bartlomiej Kaminski Manuel de la Rocha	August 2003	F. Sy 89750
WPS3109	The Impact of China's WTO Accession on East Asia	Elena Ianchovichina Terri Walmsley	August 2003	S. Lipscomb 87266
WPS3110	Governance of Public Pension Funds: Lessons from Corporate Governance and International Evidence	David Hess Gregorio Impavido	August 2003	P. Braxton 32720
WPS3111	International Trade and Wage Discrimination: Evidence from East Asia	Günseli Berik Yana van der Meulen Rodgers Joseph E. Zveglich, Jr.	August 2003	M. Ponglumjeak 31060